

# Data Aggregation Approach Using Neural Network in Wireless Sensor Networks

Amanpreet Kaur<sup>1</sup>, Vinni Kaur<sup>2</sup>, Arvind Pathak<sup>3</sup>  
Electronics and Communication Engineering<sup>1,2,3</sup>  
Lingaya's University<sup>1,2,3</sup>  
Faridabad (NCR), India<sup>1,2,3</sup>

**Abstract**— A wireless sensor network is one of the busiest networks because of multicast and broadcast network. In case of separate communication it gives high energy loss because there is requirement of some mechanism that can select multiple communications in single communication. This kind of merging is called data aggregation. Aggregation process is itself efficient process that uses the improved greedy approach. Here we have presented a secure and authentication approach for data aggregation. The user authentication is performed using secure hash algorithm. It will reflects participating node over the communication. Final stage is proposed work of neural network for checking bad packet communication over the network. The derived result showed the presented work which is more reliable and efficient than existing approach

**Keywords**-WSN; sensor node; Data aggregation; secure hash algorithm; neural network.

## I. INTRODUCTION

Recent advances in micro-electro-mechanical systems (MEMS) technology, wireless communications, and digital electronics have enabled the development of low-cost, low-power, multifunctional sensor nodes that are small in size and communicate unfettered in short distances [1]. These tiny sensor nodes, which consist of sensing, data processing, and communicating components, leverage the idea of sensor networks based on collaborative effort of a large number of nodes. Sensor networks are collection of sensor nodes which co-operatively send sensed data to base station. As sensor nodes are battery driven, an efficient utilization of power is essential in use networks for long duration hence it is needed to reduce data traffic inside sensor networks, reduce amount of data that need to send to base station. Sensor networks represent a significant improvement over traditional sensors, which are deployed in the following two ways: Sensors can be positioned far from the actual phenomenon, i.e., something known by sense perception. In this approach, large sensors that use some complex techniques to distinguish the targets from environmental noise are required. Several sensors that perform only sensing can be deployed. The positions of the sensors and communications topology are carefully engineered. They transmit time series of the sensed phenomenon to the central nodes where computations are performed and data are fused.

## II. WIRELESS SENSOR NETWORK

A wireless sensor network (WSN) consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants and to cooperatively pass their data through the network to a main location. The more modern networks are bi-directional, enabling also to control the activity of the sensors. The development of wireless sensor networks was motivated by military applications such as battlefield surveillance; today such networks are used in many industrial and consumer application, such as industrial process monitoring and control, machine health monitoring, and so on.

In computer science and telecommunications, wireless sensor networks are an active research area with numerous workshops and conferences arranged each year. Sensors integrated into structures, machinery, and the environment, coupled with the efficient delivery of sensed information, could provide tremendous benefits to society.

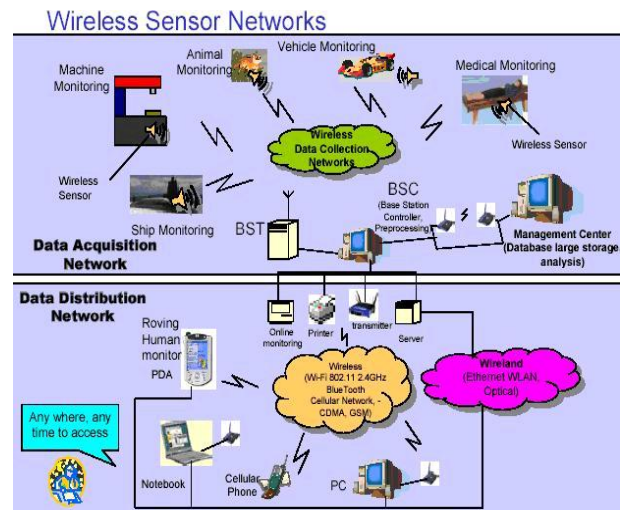


Figure 1. A typical Wireless Sensor Networks

## III. DATA AGGREGATION

Data aggregation techniques explore how the data is to be routed in the network as well as the processing method that are applied on the packets received by a node. They have a great impact on the energy consumption of nodes and thus on network efficiency by reducing number of transmission or

length of packet. Elena Fosolo et al in [3] defines the in-network aggregation process as follows: "In-network aggregation is the global process of gathering and routing information through a multi-hop network, processing data at intermediate nodes with the objective of reducing resource consumption (in particular energy), thereby increasing network lifetime."

#### IV. SECURE HASH ALGORITHM

SHA-1 produces a 160-bit message digest based on principles similar to those used by Ronald L. Rivest of MIT in the design of the MD4 and MD5 message digest algorithms, but has a more conservative design.

The original specification of the algorithm was published in 1993 as the Secure Hash Standard, FIPS PUB 180, by US government standards agency NIST (National Institute of Standards and Technology). This version is now often referred to as SHA-0. It was withdrawn by NSA shortly after publication and was superseded by the revised version, published in 1995 in FIPS PUB 180-1 and commonly referred to as SHA-1. SHA-1 differs from SHA-0 only by a single bitwise rotation in the message schedule of its compression function; this was done, according to NSA, to correct a flaw in the original algorithm which reduced its cryptographic security. However, NSA did not provide any further explanation or identify the flaw that was corrected. Weaknesses have subsequently been reported in both SHA and SHA-1. SHA-1 appears to provide greater resistance to attacks, supporting the NSA's assertion that the change increased the security.

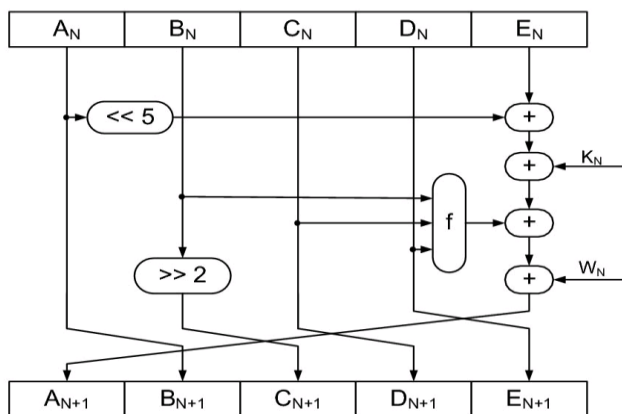


Figure 2. The secure hash algorithm (SHA-1)

#### Algorithm

The basic SHA-1 algorithm is presented as follows:

1) The algorithm starts off by initializing the five sub-registers of the first 160-bit register X labeled  $H_0$ ,  $H_1$ ,  $H_2$ ,  $H_3$ , and  $H_4$  as follows:

$H_0=67452301$ ;  $H_1=EFCDAB89$ ;  $H_2=98BADCFE$ ;  
 $H_3=10325476$ ;  $H_4=C3D2E1F0$ ;

2) From here onwards, SHA-1 iterates through each of the 512-bit message blocks viz.  $m_0, m_1, m_2, \dots, m_{n-1}$ . For each of the message block, do the following:

Write  $m_j$  as a sequence of sixteen 32-bit words,

$m_j = W_0 \parallel W_1 \parallel W_2 \parallel \dots \parallel W_{15}$

Compute the remaining sixty four 2-bit words as follows:

$W_t = (W_{t-3} \text{ xor } W_{t-8} \text{ xor } W_{t-14} \text{ xor } W_{t-16})$

Cyclic shift of  $W_t$  by 1 i.e.  $S^1(W_t)$

Copy the first 160 bit register into the second register as follows:

$A = H_0$ ;  $B = H_1$ ;  $C = H_2$ ;  $D = H_3$ ;

$E = H_4$ ;

This step involves a sequence of four rounds, corresponding to four intervals  $0 \leq t \leq 19$ ,  $20 \leq t \leq 39$ ,  $40 \leq t \leq 59$ ,  $60 \leq t \leq 79$ . Each round takes as input the current value of register X and the blocks  $W_t$  for that interval and operates upon them for 20 iterations as follows:

For  $t = 0$  to 79,

$T = S^5(A) + f_t(B, C, D) + E + W_t + K_t$

$E = D$ ;  $D = C$ ;  $C = S^{30}(B)$ ;

$B = A$ ;  $A = T$

Once all four rounds of operations are completed, the second 160-bit register ( $A, B, C, D, E$ ) is added to the first 160-bit register ( $H_0, H_1, H_2, H_3, H_4$ ) as follows:

$H_0 = H_0 + A$ ;

$H_1 = H_1 + B$ ;

$H_2 = H_2 + C$ ;

$H_3 = H_3 + D$ ;

$H_4 = H_4 + E$ ;

3) Once the algorithm has processed all of the 512-bit blocks, the final output of X becomes the 160-bit message digest.

#### V. NEURAL NETWORK:

The term neural network was traditionally used to refer to a network or circuit biological neurons. The modern usage of the term often refers to artificial neural networks, which are composed of artificial neurons or nodes. The input to the NN is chosen to include previous output samples of the modeling sensor node and the current and previous output samples of neighboring sensors. The model is based on a new structure of a back propagation-type NN.

##### A. Back Propagation Neural Network:

Back propagation is a common method of teaching artificial neural networks how to perform a given task. It was first described by Arthur E. Bryson and Yu-Chi Ho in 1969, but it wasn't until 1974 and later, through the work of Paul Werbos, David E. Rumelhart, Geoffrey E. Hinton and Ronald J. Williams, that gained recognition, and it led to a "renaissance" in the field of artificial neural network research. It is a supervised learning method, and is a generalization of the delta rule. It requires a teacher that knows, or can calculate, the desired output for any input in the training set. It is most useful for feed-forward networks (networks that have no feedback, or simply, that have no connections that loop). The term is an abbreviation for "backward propagation of errors". Back propagation requires that the activation function used by the artificial neurons (or "nodes") be differentiable.

##### B. Propagation:

Each propagation involves the following steps:

- Forward propagation of a training pattern's input through the neural network in order to generate the propagation's output activations.
- Back propagation of the propagation's output activations through the neural network using the training pattern's target in order to generate the deltas of all output and hidden neurons.

### C. Weight update:

For each weight-synapse:

- Multiply its output delta and input activation to get the gradient of the weight.
- Bring the weight in the opposite direction of the gradient by subtracting a ratio of it from the weight.

This ratio influences the speed and quality of learning; it is called the *learning rate*. The sign of the gradient of a weight indicates where the error is increasing; this is why the weight must be updated in the opposite direction.

### ALGORITHM

- Step 1: Initialize the weights in the network
- Step 2: Train the input for N datasets
- Step 3: For Each Training Set I in N  
[Repeat Steps 4 to ]
- Step 4: Set O= NaturalOutput(Network, I)
- Step 5: Set T= Teacher Output of I
- Step 6: Calculate Error T-O
- Step 7: Compute W1=All Weights from Hidden Layer to Output Layer
- Step 8: Compute W2=All Weights from Input Layer to Hidden Layer
- Step 9: Update Weights in the Network
- Step 10: Go to Step 3 Until Classification not done correctly.

### VI. MATLAB

MATLAB (matrix laboratory) is a numerical computing environment and fourth-generation programming language. Developed by Math Works, MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages, including C, C++, Java, and FORTRAN. It also provide interactive environment that enables you to perform computationally intensive tasks faster than with traditional programming languages such as C, C++, and FORTRAN.

Although MATLAB is intended primarily for numerical computing, an optional toolbox uses the MuPADsymbolic engine, allowing access to symbolic computing capabilities. An additional package, Simulink, adds graphical multi-domain simulation and Model-Based Design for dynamic and embedded systems

### VII. RESULT

We begin the task with modeling a Wireless Sensor Network, comprising of ten nodes initially. The scenario

containing the sensor nodes along with a base station and router is developed.

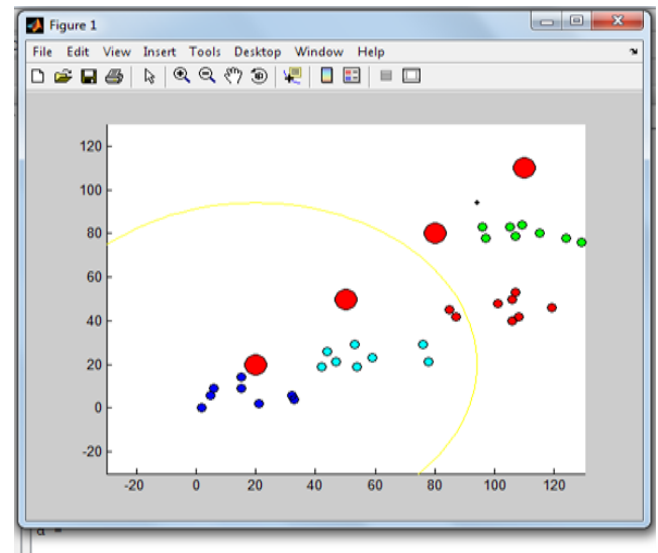


Figure 3. The modeled Wireless Sensor Network using Matlab

The figure above represents the nodes thus labeled and placed in the Wireless Sensor Networks. The various nodes are labeled as node 1 till node 10. The position of node1 is fixed and rests of the nodes are placed in the network using the "flood" function available with the Matlab. The node is red color represent the sink node of the network

The sensor nodes present in the modelled wireless sensor network use the Secure Hash Algorithm (SHA-1) for the transmission of data from the sink to other nodes and other nodes to the sink node.

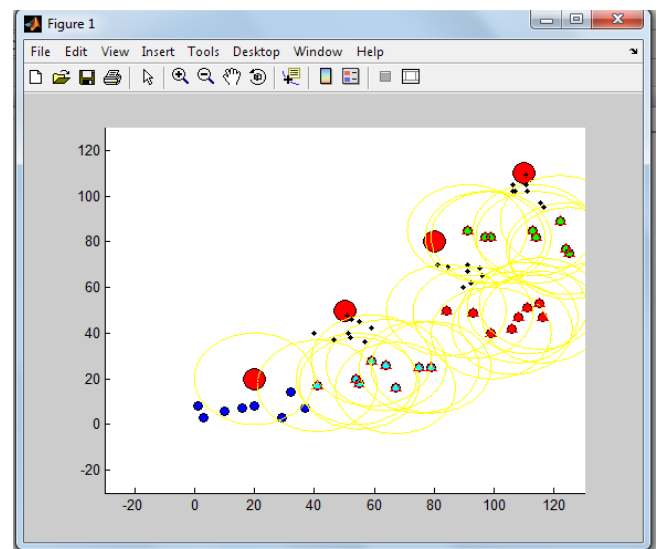


Figure 4. Data flow among the node in WSN.

Once the data transmission is authenticated by implementing SHA-1 algorithm, a dialog box prompts the user ensuring the authenticity of the data being transmitted among the sink node and other nodes and vice versa. The keys for transmission are generated and distributed. The neural network tool box is used to train the ANN-WSN network.

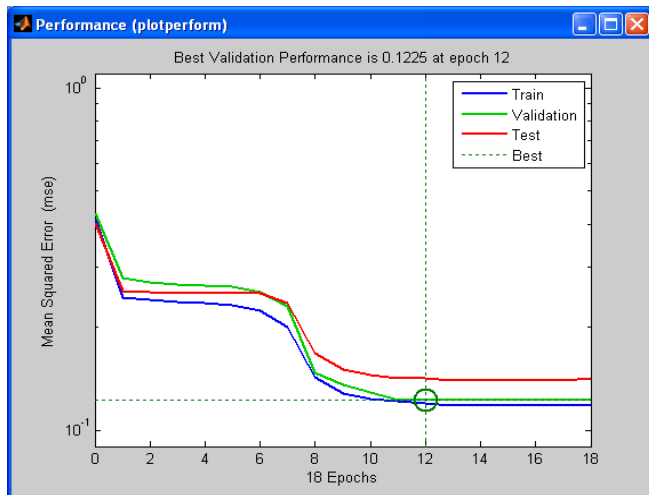


Figure 5. Performance plot.

The result is reasonable because of the following considerations:

- The final mean-square error is small.
- The test set error and the validation set error has similar characteristics.
- No significant over fitting has occurred by iteration 12 (where the best validation performance occurs).
- The training phase is related to the validation phase, although the testing phase is different from the other two phases shown during the learning process of the system. The validation and test curves are very similar.

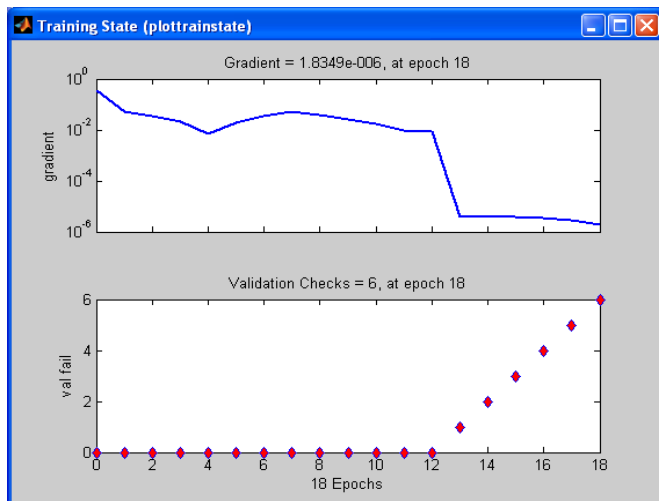


Figure 6. Training state.

The above graph leads us to the conclusion that the neural network thus established starts learning at the 12<sup>th</sup> epochs. The number of validation checks being performed are 6 at epoch 18.

The graph above mainly contains two graphs. One of the graphs is plotted taking Val fail value and number of epochs. In this graph The ANN-WSN thus developed didn't perform any learning until the simulation reach epochs number 12. After the 12 epoch, the neural network starts learning

linearly throughout the end of the simulation. The other graph showed the gradient variation along with number of epochs. The graph showed a s-shaped curve to the gradient value of 1.8349e006. Thus indicating the variation along the s curve shown in the training state.

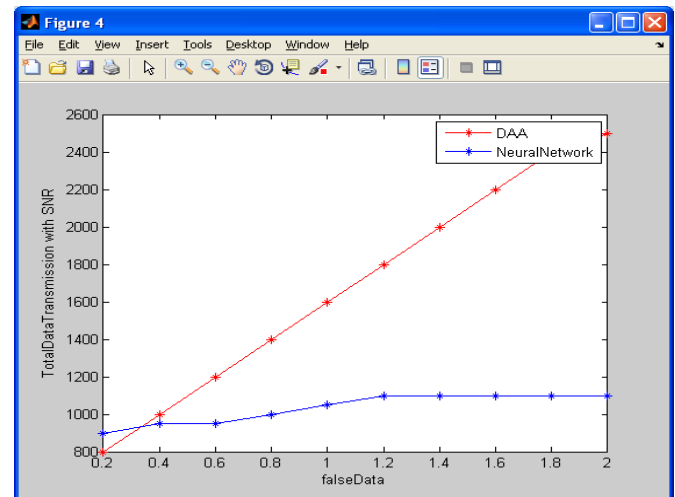


Figure 7. Graph showing total data transmission with SNR.

The above shown graph thus obtained on training the ANN-WSN. The graph plots Total Transmission with SNR (Signal to Noise Ratio). The red colored line in the graph indicate the DAA (Data Aggregation and Authentication) protocol performance whereas the blue coloured line indicate the performance of neural network based network. It can be concluded from the plot that Neural network reduce the false data detection better in comparison to the DAA algorithm

## VIII. CONCLUSION

The similarities between wireless sensor networks and neural networks suggest that combining these technologies is possible. In this dissertation work we describe our experiences with constructing a combined system in order to resolve some of the long standing issues in wireless sensor network technology. Our results show that such a combined system can indeed be built, and that its performance is on par with traditional systems.

Due to the unique property of sensor networks, public keys do not need to be authenticated in the same way as it is done in the Internet environment (i.e., using certificates); instead, public keys can be authenticated using one-way hash functions (SHA-1), which are much more efficient than signature verification on certificates can increase efficiency of data transfer with reduction of data losses and delays.

## ACKNOWLEDGMENT

I am highly thankful to my guide Mr. Arvind Pathak, Assistant Professor, Electronics and Communication Engineering Department of Lingaya's University, Faridabad, for his advice, motivation, guidance, moral support, efforts and the attitude with which he solved all of my queries in making this thesis possible. It has been a great honor to work under him.

I am also thankful to Asst. Prof. Pragati Kapoor, Head of Department, Electronics and Communication Engineering Department of Lingaya's University, Faridabad, for providing us with adequate infrastructure in carrying the research work.

#### REFERENCES

- [1] F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cyirci, Wireless sensor networks: A survey, Computer Networks, Vol. 38, No. 4, Page(s): 393-422, 2002.
- [2] "Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems" by Mohammad Ilyas and Imad Mahgoub
- [3] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-Network Aggregation Techniques for Wireless Sensor Networks: A Survey", IEEE Wireless communication 2007.
- [4] K. Kalpakis, K. Dasgupta and P. Namjoshi, "Maximum Lifetime Data Gathering and Aggregation in Wireless Sensor Networks", in Proceedings of IEEE Networks 2002
- [5] M. Cardei and D. Du. Improving wireless sensor network lifetime through power aware organization. ACM Journal of Wireless Networks, 11(3):333{340, May 2005.
- [6] L. Fausett, *Fundamentals of Neural Networks*. Prentice Hall, 1994.